

• 药剂与工艺 •

近红外光谱快速预测天舒片包衣终点研究

夏春燕^{1,2}, 徐芳芳^{2,3,4,5}, 张欣^{2,3,4,5}, 徐冰⁶, 包乐伟^{2,3,4,5}, 王振中^{2,3,4,5}, 肖伟^{1,2,3,4,5*}

1. 南京中医药大学, 江苏 南京 210023

2. 江苏康缘药业股份有限公司, 江苏 连云港 222001

3. 中药制药过程新技术国家重点实验室, 江苏 连云港 222001

4. 中成药智能制造国家地方联合工程研究中心, 江苏 连云港 222001

5. 中药提取精制新技术重点研究室, 江苏 连云港 222001

6. 北京中医药大学 中药信息学系, 北京 100029

摘要: **目的** 应用近红外光谱(NIRS)技术快速预测天舒片薄膜包衣终点。**方法** 首先,采用组合间隔偏最小二乘法(synergy interval partial least square, siPLS)优选建模区间,并对预处理方法进行筛选;其次,根据优选的建模条件,基于建立的参考光谱库,采用一致性检验和相似度匹配建立校正模型;随后,通过对比测试光谱与参考光谱间的相似度大小,设定阈值,确定相似度、阈值与包衣终点的关系;最后,采用验证集光谱对模型性能进行验证。**结果** 在包衣过程中,当连续多条测试光谱的相似度超过阈值时,指示为包衣终点。**结论** 一致性检验和相似度匹配模型性能良好,能够敏锐地反映包衣薄膜的均匀性,并准确预测包衣终点。该研究对减小包衣的批间差异性和降低原料损耗,提高中药固体制剂的包衣生产效率和提升质量均一性具有重要意义。

关键词: 天舒片; 近红外光谱技术; 包衣; 一致性检验; 相似度匹配; 组合间隔偏最小二乘法; 质量均一性

中图分类号: R283.6 文献标志码: A 文章编号: 0253-2670(2019)21-5223-08

DOI: 10.7501/j.issn.0253-2670.2019.21.014

Study on fast prediction of film coating terminal point of Tianshu Tablets by near-infrared spectroscopy

XIA Chun-yan^{1,2}, XU Fang-fang^{2,3,4,5}, ZHANG Xin^{2,3,4,5}, XU Bing⁶, BAO Le-wei^{2,3,4,5}, WANG Zhen-zhong^{2,3,4,5}, XIAO Wei^{1,2,3,4,5}

1. Nanjing University of Chinese Medicine, Nanjing 210023, China

2. Jiangsu Kanion Pharmaceutical Co., Ltd., Lianyungang 222001, China

3. State Key Laboratory of New-tech for Chinese Medicine Pharmaceutical Process, Lianyungang 222001, China

4. National & Local Joint Engineering Research Center on Intelligent manufacturing of TCM, Lianyungang 222001, China

5. Key Laboratory of New Technology for Extraction and Refining of Traditional Chinese Medicine, Lianyungang 222001, China

6. Department of Traditional Chinese Medicine Informatics, Beijing University of Chinese Medicine, Beijing 100029, China

Abstract: Objective To investigate how to rapidly predict coating terminal point in the thin film coating of Tianshu Tablets by near-infrared spectroscopy (NIRS). **Methods** Firstly, synergy interval partial least square (siPLS) was used to optimize the modeling intervals and the pretreatment methods were screened. Secondly, according to the optimal modeling parameters, the conformity test and similarity matching were used to establish calibration models based on the established reference spectra library. Subsequently, by comparing the similarities between the test spectra and the reference spectra, the threshold value was set. The relation among the similarities, threshold value and the coating end point was obtained. Finally, the performance of the calibration model was verified by

收稿日期: 2019-05-20

基金项目: 国家“重大新药创制”科技重大专项“国家工信部智能制造综合标准化与新模式应用项目”(2018ZX09201010)

作者简介: 夏春燕(1994—),女,在读硕士,研究方向为新型中药制剂的研究与开发。Tel: (0518)81152363 E-mail: xiachytz46@163.com

*通信作者 肖伟,研究员级高级工程师,博士,研究方向为中药新药的研究与开发。Tel: (025)87181851 E-mail: kanionlunwen@163.com

the validation set. **Results** During the coating process, the similarities between the test spectra and the reference spectra were gradually increasing until the similarities of continuous test spectra exceeded the threshold, thus, indicating the end point. **Conclusion** Conformity test and similarity matching can sensitively monitor the change of similarities among spectra and the corresponding trend. Both models have high performance, which can accurately predict the coating terminal point. It's of great significance to reduce batch-to-batch variation in film coating and the loss of coating materials, and improve the coating efficiency and the quality uniformity of solid preparation.

Key words: Tianshu Tablets; near-infrared spectroscopy; film coating; conformity test; similarity matching analysis; synergy interval partial least squares (siPLS); quality uniformity

天舒片由川芎和天麻 2 味药组成, 临床用于治疗神经性偏头痛及高血压性头痛^[1]。为了使片剂免受光致分解或酸降解等外部影响, 掩盖味道, 或控制药效成分的定点定时释放, 会对片剂进行包衣。包衣是一个多单元多变量的操作过程, 由于包材喷涂、分布和干燥同时进行, 许多关键工艺参数 (critical process parameters, CPPs) 都会影响片剂包衣的生长和均匀性。喷枪速率的变化会显著影响片剂到达一定厚度所需的时间^[2]。进出气的温度和喷枪速率是影响高分子聚合物成膜的重要因素。当达到最低成膜温度时, 需确保片剂有适当的含水率^[3]。在工业生产时确保包衣厚度和包衣质量的均一性显得至关重要。包衣厚度的测量并非如想象中那般容易。研究表明, 片剂中心和侧面的包衣厚度是存在显著性差异的^[4]。包衣时, 片心粗糙的表面在主包衣层形成前就已完成了填充^[5], 因此, 根据包衣前后的片剂平均质量计算厚度的预测值相比直接测量的预测值会更准确些^[6]。然而, 这些方法均在少量样品上重复测量, 且耗时耗力。因此更倾向于一些适合大量样品的快捷的研究方法, 如 X 射线计算机断层技术^[4]、声显微技术^[7]。包衣过程中, 一些聚合物溶液是直接喷雾干燥的, 并非在片剂表面形成包衣层, 这可能导致不均匀包衣层^[8]。当出现这种情况时, 通常选择继续包衣, 延迟包衣终点, 最终包衣过厚, 影响了药物的溶出。

一致性检验 (conformity test) 和相似度匹配 (similarity match) 属于定性研究方法。一致性检验主要用于药品的真伪鉴别^[9-10], 厂家、产地的归属鉴定^[11-12]及药品供应链全程质量监控^[13-14]及生物等效性分析^[15]等方面。相似度匹配则广泛应用于药材厂家或产地鉴别^[16-17]、真伪判断及药材质量控制^[18]研究。目前鲜有将这 2 种方法应用于片剂薄膜包衣领域的研究报道。本研究以天舒片为研究对象, 收集参考样品 (包衣终点样品) 和测试样品 (不同包衣时间的样品), 采集近红外漫反射光谱; 并采用一

致性检验和相似度匹配分析方法建立校正模型。以期该方法在包衣过程中的应用能够提升片剂包衣薄膜的均匀性, 实现准确的终点判别。

1 仪器与材料

1.1 仪器

BGB-150 C 型高效包衣机, 宜春万申制药机械有限公司; Bruker MPA 型傅里叶变换近红外光谱仪, 配积分球漫反射检测器, 德国布鲁克光学仪器公司。

1.2 材料

天舒片由江苏康缘药业股份有限公司提供, 其片心为直径 10 mm、厚度 4 mm 的双凸片; 蓝色胃溶型包衣薄膜预混剂, 批号 F1810328, 购自上海新菲尔物科技有限公司。

2 方法

2.1 包衣过程和取样方法

1 批天舒片心分成 4 份 (110、110、100、100 kg) 进行薄膜包衣, 随机取一份天舒片心投入包衣锅内。包衣开始后, 每隔 30 min 从设备预留取样口迅速取样, 每次取约 100 片。取样后样品用自封袋密闭保存。本研究共完成 9 个批次的取样, 制剂批号分别为 180809、180901、180902、180903、180904、180905、180906、180907、180908, 见表 1。本次共收集了 68 个时间点的过程样品和 9 个包衣终点的样品^[19]。

2.2 近红外光谱采集

近红外光谱仪波数为 10 000~4 000 cm^{-1} , 分辨率为 16 cm^{-1} , 以空气为参比扣除背景, 每条 NIRS 为 64 次扫描的平均值。取片心和每个时间点的包衣片 10 片进行近红外漫反射光谱采集。将手持式近红外探头贴紧药片凸面, 使近红外光束集中在凸面的中心位置, 每片正反面分别重复扫描 3 次, 每片扫描后获得 6 条 (2 面×3 次/面) NIRS。

2.3 参考光谱库的建立

以批次 180809、180901、180902、180903、

表 1 不同批次包衣过程样品取样时间点

Table 1 Sampling time of samples

批号	取样时间点/min	终点时间/min
180809	0、30、60、90、120、150、180、210	220
180901	20、50、80、110、140、170、200	230
180902	30、60、90、120、150、180	220
180903	10、40、70、100、130、160、190、220	230
180904	0、30、60、90、120、150、180、210	220
180905	0、30、60、90、120、150、180、210	220
180906	10、40、70、100、130、160、190、220	230
180907	0、30、60、90、120、150、180、210	220
180908	30、60、90、120、150、180、210	220

180904、180905、180906、180907、180908 的包衣终点样品的 NIRS 建立参考光谱库。

2.4 样品集划分

按制剂批次顺序, 以批次 180809、180901、180902、180903、180904、180905 共 45 个时间点的过程样品为校正集, 以批次 180906、180907、180908 共 23 个时间点的过程样品为验证集。

2.5 一致性检验

一致性检验^[10]可用于快速比较测试光谱与参考光谱间的相似性, 其原理为计算参考光谱集每个波长点的平均值和标准偏差 (S), 以波长点的平均值加减一定倍数的标准偏差作为该波长点的控制区间; 如测试光谱每个波长点的吸光度没有超出预定的控制范围, 则说明测试样品与参考样品具有相似性。控制区间计算公式为参考光谱吸光度平均值 $\pm$ CI 限度值 $\times S$, 其中 CI (conformity index) 表示一致性指数, 可通过以下步骤确定: (1) 计算参考光谱每个波长点 i 处吸光度的平均值 $A_{i,参}$ 和标准偏差 $S_{i,参}$; (2) 计算样品每个波长点处的 CI_i 值, 计算公式: $CI_i = (A_{i,样} - A_{i,参}) / S_{i,参}$, 其中, $A_{i,样}$ 为样品测试光谱在波长点 i 处吸光度; (3) 样品的 $CI_{样}$ 值为该样品在每个波长点处的 CI_i 值的最大值, 即 $CI_{样} = \max(CI_{i,样})$, 同样的, 参考光谱的 $CI_{参}$ 值为该参考样品在每个波长点处的 CI_i 值的最大值, 即 $CI_{参} = \max(CI_{i,参})$ 。

在建模阶段, 通过比较合格光谱和不合格光谱的 CI 值, 选择合适的 CI 限度值。而在应用阶段可直接将测试光谱的 CI 值与之前设定的 CI 限度进行比较, 从而快速简单地判断。

2.6 相似度匹配

相似度匹配是一种光谱分类技术, 通过比较测试光谱与参考光谱在波段范围内的光谱信息得到两者间的匹配程度。相似度匹配是基于 Gram-Schmidt 矢量正交方法整体评价相似度^[20-21], 基本原理如下。

设参考图谱库是包含 m 条光谱且波长点数为 n 的矩阵 X , 将其正交分解为得分矩阵 T 和载荷矩阵 P , 其中 r 为矩阵 X 的秩, 见式 (1)。

$$X = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1r} \\ t_{21} & t_{22} & \cdots & t_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ t_{m1} & t_{m2} & \cdots & t_{mr} \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ p_{r1} & p_{r2} & \cdots & p_{rn} \end{bmatrix} \quad (1)$$

其中 $p_j = [p_{j1}, p_{j2}, \cdots, p_{jn}]$ ($j=1, 2, \cdots, r$) 为一组相互正交的向量。由于 p 由正交向量组成, 所以 $T = XP^T$ 。测试光谱数据为 x_{new} , 可利用式 (2) 计算得分向量 t_{new} 。

$$t_{new} = x_{new} P^T = [s_1, s_2, \cdots, s_n] \quad (2)$$

$j=1$ 时, 根据 $e_{i1} = |s_1 - t_{i1}|^2$ ($i=1, 2, \cdots, m$) 计算得到 $e_1 = [e_{11}, e_{21}, e_{31}, \cdots, e_{m1}]$, 并得到向量 e_1 中值最小的元素 e_{k1} , 对应的 t_{k1} 记为 t_1' , 接着, 相应地计算其余光谱的 $t_2', t_3', \cdots, t_r'$, 得到测试光谱与参考光谱间的残差光谱 e , 见式 (3), 该残差光谱反映了测试光谱与参考光谱间的差异。

$$e = [s_1 - t_1', s_2 - t_2', \cdots, s_r - t_r'] \begin{bmatrix} p_1 \\ p_1 \\ \cdots \\ p_r \end{bmatrix} \quad (3)$$

计算测试光谱与参考光谱之间的相似度匹配值 (similarity match value, SMV), 见公式 (4)。

$$SMV = (1 - \|e\| / \|x_{new}\|) \times 100 \quad (4)$$

e 为测试样品的残差光谱, x_{new} 为测试光谱数据。SMV 反映了测试光谱的残差平方和。SMV 分布在 0~100, 越接近 100, 说明测试样品光谱与参考光谱信息越接近。

2.7 组合间隔偏最小二乘法^[22-23]

组合间隔偏最小二乘法 (synergy interval partial least square, siPLS) 是对间隔偏最小二乘法 (interval partial least square, iPLS) 的扩展, 能够弥补 iPLS 一个子区间建模产生的不足。siPLS 建模的一般步骤如下: (1) 将全光谱等分为 N 个子区间; (2) 将 N 个子区间排列组合成不同的子区间 (一般为 2、3 或 4 个子区间进行组合), 并用于建立偏最小二乘模型; (3) 根据交叉验证均方根误差 (root mean square

error of cross validation, RMSECV) 选择精度最高的模型, 并选择相应的组合区间, RMSECV 最小时的模型的精度最高; (4) 优化所选组合区间, 以步骤 3 中选择的区间为中心, 对波长区域进行单向或双向扩展, 得到一个最佳波长区间; (5) 最后将具有高精度模型的区间进行建模预测, 并根据最小 RMSECV 确定最优组合区间。

2.8 模型评价参数

基于所建的参考光谱库, 以校正集样品的 NIRS 为测试光谱, 建立定性校正模型。根据设定的阈值, 以模型预测的正确率 (W)^[16-17] 为指标优化建模参数, 考察模型性能, W 的计算方法见公式 (5)。其中, n 为正确预测的样品数, N 为样品总数。

$$W = n/N \quad (5)$$

模型预测标准: 当样品的光谱相似度小于相似度阈值时, 模型预测为过程样品; 当光谱相似度均大于或等于相似度阈值时, 模型预测其为包衣终点的样品。

2.9 数据分析

一致性检验在 OPUS 化学计量学软件 (version 7.5, 德国布鲁克公司) 上完成; 相似度匹配在 TQ Analyst 光谱分析软件 (version 9.0, 美国 Thermo 公司) 上完成; 波段筛选在 MATLAB 软件 (version 2016 b, 美国 MathWorks 公司) 上完成。

3 结果

3.1 一致性检验

3.1.1 波段选择 NIRS 主要是分子倍频和合频的吸收, 除了包含片心和高分子薄膜的化学信息外, 还包含大量重叠信息和噪音, 采用全光谱建模时, 模型的预测精度未必能达到最佳值。因此, 需要剔除冗余信息, 提高分析模型的预测精度和建模效率。

基于所建的参考光谱库, 以校正集原始光谱为测试光谱, 光谱范围为 $10\,000 \sim 4\,000\text{ cm}^{-1}$ 的建模参数建立模型, 得到校正集光谱的 CI 值。以每个时间点样品的 NIRS 吸光度为自变量, 对应光谱的 CI 值为因变量, 构建 45 个样品光谱组成的数据集建立 siPLS 模型。将近红外全光谱等分为 20 个子区间, 采用 3 个子区间组合的形式筛选波段。所得区间为 $7\,290.06 \sim 6\,996.91\text{ cm}^{-1}$ 、 $5\,785.76 \sim 5\,477.19\text{ cm}^{-1}$ 、 $4\,597.75 \sim 4\,312.32\text{ cm}^{-1}$ 。从片心、包衣过程样品和终点样品光谱图 (图 1) 上可以看出, 包衣过程光谱和终点光谱在 $7\,290.06 \sim 6\,996.91\text{ cm}^{-1}$ 和 $4\,597.75 \sim 4\,312.32\text{ cm}^{-1}$ 2 个区间内有 2 个尖峰, 而

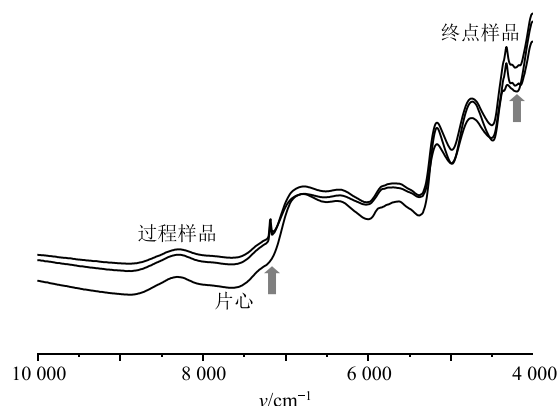


图 1 片心、过程样品和终点样品光谱图

Fig. 1 Spectra of core tablets, process tablets, and finished tablets

片心的光谱图上未出现该峰, 推测尖峰可能是包衣薄膜的特征峰。 $7\,290.06 \sim 6\,996.91\text{ cm}^{-1}$ 的差异对应 C-H 一级倍频伸缩振动基频与 C-H 变形振动基频组合频; 在 $4\,597.75 \sim 4\,312.32\text{ cm}^{-1}$ 的光谱差异属于 C-H 伸缩振动基频与 C-H 变形振动基频组合频和 C-H 二级倍频伸缩振动基频。而 $5\,785.76 \sim 5\,477.19\text{ cm}^{-1}$ 的信息在光谱图上没有表现出十分显著性的差异, 但与文献报道^[24] 的包衣辅料 $5\,986 \sim 5\,369\text{ cm}^{-1}$ 的特征吸收区相吻合, 推断为包衣薄膜的特征峰。

3.1.2 预处理方法选择 受背景等环境因素的影响及仪器性能的制约, 采集的光谱包含了噪音和失真信息, 有必要对其进行预处理, 剔除冗余信息, 强化特征吸收。基于所建立的参考光谱库, 以校正集光谱为测试光谱, 光谱范围为优选的组合区间, 分别在不同的预处理方法条件下建立模型。本研究考察了标准正态变换 (standard normal variate, SNV) 与导数法 (一阶导数、二阶导数) 结合对光谱预处理。SNV 可用来消除固体颗粒大小, 表面散射以及光程变化对漫反射光谱的影响, 导数法可以消除基线偏移和背景干扰, 分辨重叠峰, 提高分辨率^[25]。不同预处理方法的建模结果见表 2。本研究中, CI 限度值设定为所有参考光谱 CI 值的最大值, 即

表 2 不同预处理方法的比较

Table 2 Comparison of different pretreatment methods

预处理方法	设定的 CI	样品总数	正确预测数	正确率/%
SNV	3.25	45	42	93.75
一阶导数	2.67		43	95.56
一阶导数+SNV	3.30		44	97.78
二阶导数	3.60		42	93.75
二阶导数+SNV	3.81		42	93.75

$CI_{\text{限}} = \text{Max}(CI_{\text{参}})$ 。

根据一致性检验预测结果可知, 当一阶导数结合 SNV 处理光谱时, 模型预测正确率最高。此时, $CI_{\text{限}} = \text{Max}(CI_{\text{参}}) = 3.3$ 。

3.1.3 一致性检验模型建立 基于参考光谱库, 以校正集光谱为测试光谱, 在 $7\ 290.06 \sim 6\ 996.91\ \text{cm}^{-1}$ 、 $5\ 785.76 \sim 5\ 477.19\ \text{cm}^{-1}$ 、 $4\ 597.75 \sim 4\ 312.32\ \text{cm}^{-1}$ 3 个波段内, 以一阶导数+SNV 为预处理方法, $CI_{\text{限}}$ 为 3.3 的条件下进行一致性检验。

结果表明, 该模型预测校正集光谱均为包衣过程中的光谱。由图 2 可知, $CI_{\text{限}}$ 为 3.3 时, 校正集光谱和参考光谱区分明显, 从整体看, 随着包衣时间延长, 校正集光谱的 CI 值向 $CI_{\text{限}}$ 趋近, 表明校正集光谱与参考光谱的相似度越高, 越靠近包衣终点。此时, 会出现一些光谱的 CI 值在 $CI_{\text{限}}$ 以下, 主要原因是包衣质量不均匀, 仍需继续包衣, 直至光谱 CI 值均在 $CI_{\text{限}}$ 以下时, 指示为包衣终点。

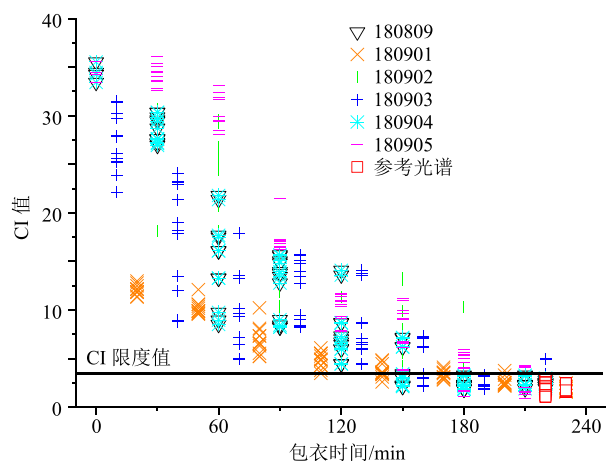


图 2 校正集光谱的 CI 值随时间变化趋势

Fig. 2 Tendency chart of CI value of calibration spectra

3.1.4 模型验证 采用验证集光谱对所建的校正模型进行验证, 结果见图 3, 模型预测正确率为 95.65%, 表明所建的校正模型可以用于天舒片包衣过程样品与终点样品间的相似度分析, 模型预测性良好, 为提升包衣质量的均一性, 准确预测包衣终点提供技术指导。

3.2 相似度匹配

3.2.1 波段选择 波段筛选的步骤如下。

首先, 从所建的参考光谱库选出 180906、180907、180908 批次的参考光谱和校正集光谱作为测试光谱, 以剩余批次的参考光谱作为参考光谱库, 光谱范围为 $10\ 000 \sim 4\ 000\ \text{cm}^{-1}$ 的建模参数建立模型, 得到每条 NIRS 的 SMV。180906、180907、180908

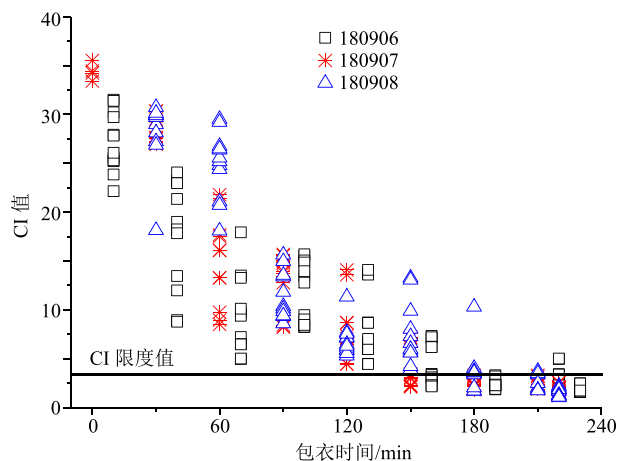


图 3 一致性检验验证模型的拟合图

Fig. 3 Fitting diagram of conformity test of validation spectra

批次参考光谱的 SMV 用于计算模型的 SMV 阈值。

其次, 以每个时间点样品的 NIRS 吸光度为自变量, 光谱所对应的 SMV 为因变量, 构建 45 个样品组成的数据集建立 siPLS 模型。将近红外全光谱等分为 20 个子区间, 联合区间数设为 3。一共得到 20 个回归模型。模型性能高低主要通过 RMSECV 判断, 选取 RMSECV 最小的区间为优选区间。优选区间为 $9\ 396.074 \sim 9\ 102.93\ \text{cm}^{-1}$ 、 $6\ 680.625 \sim 6\ 395.194\ \text{cm}^{-1}$ 、 $4\ 597.75 \sim 4\ 312.32\ \text{cm}^{-1}$ 。

其中, $9\ 396.074 \sim 9\ 102.93\ \text{cm}^{-1}$ 和 $6\ 680.625 \sim 6\ 395.194\ \text{cm}^{-1}$ 均无法反映图 1 中 2 个尖峰的信息。在此基础上, 分别采用全光谱 $10\ 000 \sim 4\ 000\ \text{cm}^{-1}$, 此次筛选区间 $9\ 396.074 \sim 9\ 102.93\ \text{cm}^{-1}$ 、 $6\ 680.625 \sim 6\ 395.194\ \text{cm}^{-1}$ 、 $4\ 597.75 \sim 4\ 312.32\ \text{cm}^{-1}$ (区间 1) 和一致性检验优选区间 $7\ 290.06 \sim 6\ 996.91\ \text{cm}^{-1}$ 、 $5\ 785.76 \sim 5\ 477.19\ \text{cm}^{-1}$ 、 $4\ 597.75 \sim 4\ 312.32\ \text{cm}^{-1}$ (区间 2) 这 3 个不同的区间建立相似度匹配模型, 考察不同组合区间对模型性能的影响, 结果见图 4, $10\ 000 \sim 4\ 000\ \text{cm}^{-1}$ 所建模型中, 校正集光谱的 SMV 随时间变化趋势更明显, 随着包衣时间的延长, 与参考光谱的匹配度越高。全光谱所反映的样品信息丰富, 能更好地呈现 SMV 的变化趋势, 模型的灵敏度高。故选择全光谱 $10\ 000 \sim 4\ 000\ \text{cm}^{-1}$ 建模。

3.2.2 预处理方法选择 分别采用 SNV、导数法 (一阶导数、二阶导数) 结合平滑法 Savitzky-Golay (SG) 滤波器平滑和 Norris Derivative (ND) 滤波器平滑进行光谱预处理, 结果见表 3。

采用 SNV 预处理光谱所建的模型无法区分校

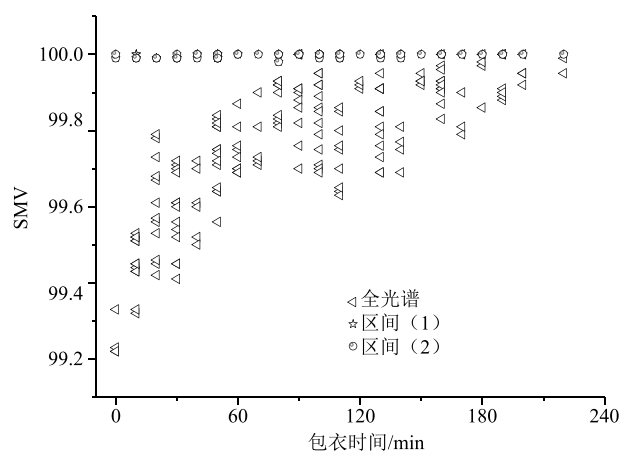


图 4 不同组合区间的比较

Fig. 4 Comparison of different intervals

表 3 不同预处理方法的比较

Table 3 Comparison of different pretreatment methods

预处理方法	阈值	正确预测数	样品总数	正确率/%
SNV	99.94	0	45	0
SNV+一阶导数	99.50	41		91.10
SNV+一阶导数+SG	99.90	43		95.56
SNV+二阶导数	98.47	45		100.00
SNV+二阶导数+SG	99.23	43		95.56
一阶导数+ND	99.96	41		91.10
二阶导数+ND	99.94	41		91.10

正集光谱和参考光谱。导数法结合平滑法处理光谱时,从 100 min 起,校正集光谱和参考光谱的 SMV 较接近,但二者无法清晰区分开,包衣终点趋势不明显。SNV+二阶导数预处理光谱时,通过 SMV 能够显著区分校正集光谱和参考光谱。180906、180907、180908 批次参考光谱的 SMV 正态检验结果表明, $P > 0.05$,表明参考光谱的 SMV 符合正态分布,设定校正模型 SMV 阈值为参考光谱 SMV 的置信下限,计算公式为 $\bar{x} - \sigma/n^{1/2}z_{\alpha/2}$ (σ 为标准差, n 为光谱数量, z 为统计量, α 为置信水平)。根据不同预处理方法所建模型的结果可知,预处理方法为 SNV+二阶导数,SMV 阈值为 98.47 时,模型预测的正确率最高。

3.2.3 相似度匹配模型建立 基于所建的参考光谱库,以校正集光谱为测试光谱,在 10 000~4 000 cm^{-1} 波段内,以 SNV+二阶导数为预处理方法,建立校正模型。模型的阈值设定为 98.47。校正模型结果表明,校正集光谱均为包衣过程光谱。如图 5 所示,校正光谱的 SMV 受时间变化影响较大,包衣

刚开始时,SMV 较小,随着包衣时间延长,SMV 逐渐增大,且越靠近阈值,表明越接近包衣终点。由于包衣机的转速、喷枪角度、喷液流量均可能影响包衣质量的均匀性,180901、180903、180904、180905 在 200 min 附近时均有部分样品的 SMV 大于阈值,这是同一批样品间和不同批次间样品包衣时均可能出现的现象^[26]。此时需要继续包衣,直到光谱的 SMV 连续大于阈值时,才达到包衣终点。

3.2.4 模型验证 将验证集光谱导入校正模型中,对模型性能进行验证。模型验证结果如图 6 所示,该模型预测验证集均为包衣过程中的样品,正确率为 100%。虽然 180906 批次在临近包衣终点时,由于包衣质量不均一,部分光谱的 SMV 大于 98.47,但是模型仍将其判定为包衣过程光谱,表明所建的相似度匹配模型能灵敏得反映包衣质量的均匀性,并且准确地预测包衣终点,可为实际生产中制剂包衣终点的判断提供技术支持。

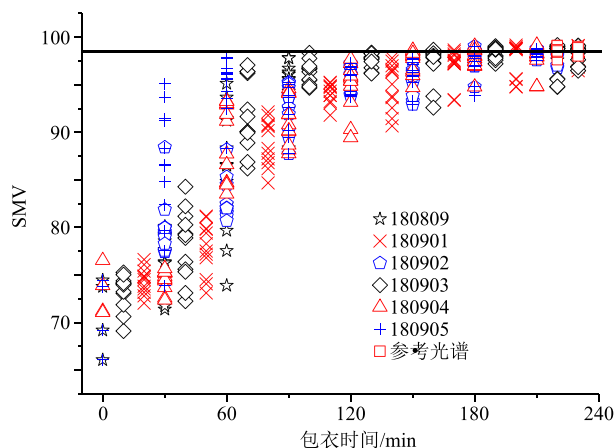


图 5 校正集光谱的 SMV 分布

Fig. 5 SMV distribution map of calibration spectra

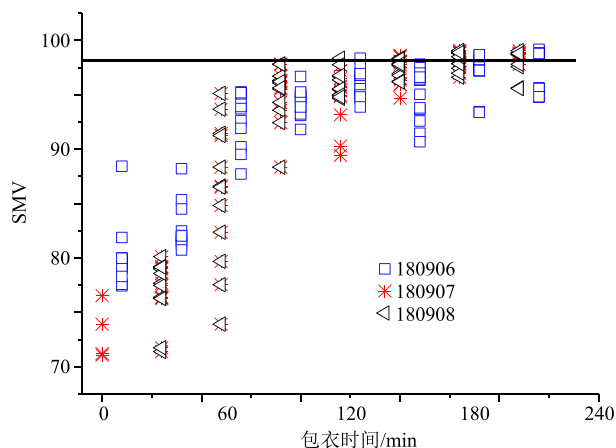


图 6 验证集光谱 SMV 分布

Fig. 6 SMV distribution map of validation spectra

4 讨论

本研究采用一致性检验和相似度匹配的分析方法建立模型预测包衣终点。结果显示,所建模型预测精度和可靠性较高,能够快速判断天舒片包衣终点和提升包衣质量均一性。本研究对检验仪器的精密度和准确度要求不高,操作简单,适用包衣过程的快速检测。

本研究按制剂批次顺序,将样品集划分为校正集和验证集。这种划分方式,能够在包衣过程中观察到测试光谱与参考光谱的相似度变化趋势,有助于提升包衣质量的均一性,准确预测包衣终点。

本研究以校正集光谱的吸光度为自变量,相应的相似度评价指标为因变量,基于 siPLS 算法优选建模区间。一致性检验的优选区间与光谱信息呈现较好的对应关系,模型性能良好。然而,在相似度匹配模型中,采用筛选区间建立的模型无法凸显测试集光谱 SMV 的变化,模型灵敏度低,这可能与 SMV 的计算原理有关。SMV 由测试光谱与参考光谱间的残差光谱计算而得。在全光谱中,残差光谱充分反映了测试光谱与参考光谱间差异的信息,模型灵敏度高。故在相似度匹配分析中采用全光谱建模。建模结果表明,采用全光谱所建的模型具有良好的预测性。

包衣并不是一个单因素生产单元,该过程受诸多因素的影响,研究的复杂性由此可见。对于功能性衣膜(具有缓控释功能)而言,例如,治疗 2 型糖尿病的固定剂量组合(fixed-dose combination, FDC)由盐酸二甲双胍的片心,格列美脲的外层包衣层和用于分隔片心和外层的中间层组成^[6]。盐酸二甲双胍具有缓释功能,格列美脲被制备成速释制剂,可见包衣膜的薄厚和均一性会影响药物的功效。促进对包衣工序的研究,并将包衣过程与药物的崩解和溶出、药效成分的含量等相结合,强化对包衣过程的理解,逐渐将包衣生产过程透明化,促进制剂工艺的发展。

参考文献

[1] 中国药典 [S]. 一部. 2015.

[2] Benet L Z, Amidon G L, Barends D M, *et al.* The use of BDDCS in classifying the permeability of marketed drugs [J]. *Pharm Res*, 2008, 25(3): 483-488.

[3] Obara S, McGinity J W. Influence of processing variables on the properties of free films prepared from aqueous polymeric dispersions by a spray technique [J]. *Int J*

Pharm, 1995, 126(1/2): 1-10.

[4] Ariyasu A, Hattori Y, Otsuka M. Non-destructive prediction of enteric coating layer thickness and drug dissolution rate by near-infrared spectroscopy and X-ray computed tomography [J]. *Int J Pharm*, 2017, 525(1): 282-290.

[5] Heinicke G, Matthews F, Schwartz J B. The effects of substrate size, surface area, and density on coat thickness of multi-particulate dosage forms [J]. *Pharm Dev Technol*, 2005, 10(1): 85-96.

[6] Hattori Y, Sugata M, Kamata H, *et al.* Real-time monitoring of the tablet-coating process by near-infrared spectroscopy-effects of coating polymer concentrations on pharmaceutical properties of tablets [J]. *J Drug Deliv Sci Technol*, 2018, 46: 111-121.

[7] Bikiaris D, Koutri I, Alexiadis D, *et al.* Real time and non-destructive analysis of tablet coating thickness using acoustic microscopy and infrared diffuse reflectance spectroscopy [J]. *Int J Pharm*, 2012, 438(1/2): 33-34.

[8] Muselik J, Dvorackova K, Krejcova K, *et al.* Pellet coating thickness determination by near-infrared reflectance spectroscopy: Comparison of two reference methods [J]. *Curr Pharm Anal*, 2010, 6(4): 225-233.

[9] 周燕, 薛磊冰, 奕水明, 等. 小通草及其伪品的近红外一致性检验和聚类分析方法研究 [J]. *中国现代应用药学*, 2017, 34(5): 697-699.

[10] 张学博, 尹利辉. 近红外光谱一致性检验方法用于快速判断药品质量的研究 [J]. *药物分析杂志*, 2011, 31(3): 603-608.

[11] Bai Y, Wang X, Song R, *et al.* Rapid identification of tiegun yam by near infrared spectroscopy combined with conformity test [A] // International Conference on Biomedical Engineering & Biotechnology (iCBEB) [C]. Macao: IEEE Computer Society, 2012.

[12] 曲佳, 王秋珍, 王杰. 近红外光谱法快速鉴别不同厂家的天王补心丸(小蜜丸) [J]. *药物评价研究*, 2014, 37(4): 346-348.

[13] 谭家群, 秦黎明, 张锐. 近红外光谱一致性检验方法在艾滋病防治药品供应链全程质量监控中的应用 [J]. *药物分析杂志*, 2012, 32(8): 1479-1481.

[14] 沙孟晨, 张海珠, 何琴, 等. 基于微量量热法的连花清瘟胶囊质量一致性评价方法的建立 [J]. *中草药*, 2017, 48(11): 2202-2206.

[15] 刘倩, 南楠, 许鸣镛, 等. FDA 对咀嚼片生物等效性要求的分析 [J]. *现代药物与临床*, 2018, 33(3): 696-702.

[16] Li W L, Han H F, Zhang L, *et al.* Manufacturer identification and storage time determination of “Dong’e

- Ejiao" using near infrared spectroscopy and chemometrics [J]. *J Zhejiang Univ-Sci B*, 2016, 17(5): 382-390.
- [17] 瞿海斌, 杨海雷, 程翼宇. 近红外漫反射光谱法快速无损鉴别阿胶真伪 [J]. *光谱学与光谱分析*, 2006, 26(1): 60-62.
- [18] 刘绍勇, 薛东升, 潘建超, 等. 基于近红外光谱技术的连翘药材质量控制方法研究 [J]. *药物分析杂志*, 2014, 34(4): 681-685.
- [19] 夏春燕, 徐芳芳, 张欣, 等. 近红外光谱分析技术快速测定天舒片的包衣薄膜厚度 [J/OL]. *中国实验方剂学志*: 1-8 [2019-08-23]. <https://doi.org/10.13422/j.cnki.syfjx.20192251>.
- [20] Ake B. Numerics of Gram-Schmidt orthogonalization [J]. *Linear Algebra Appl*, 1994, 197/198: 297-316.
- [21] 潘建超, 薛东升. 近红外光谱相似度用于中药提取物质量控制方法的研究 [J]. *中国医药工业杂志*, 2011, 42(5): 340-343.
- [22] Li W, Cheng Z, Wang Y, *et al*. Quality control of *Lonicerae Japonicae Flos* using near infrared spectroscopy and chemometrics [J]. *J Pharm Biomed Anal*, 2013, 72(4): 33-39.
- [23] 杜晨朝, 赵安邦, 吴志生, 等. 近红外光谱结合不同变量筛选方法用于金银花提取过程中绿原酸量的在线监测 [J]. *中草药*, 2017, 48(16): 3317-3321.
- [24] Möltgen C V, Puchert T, Menezes J C, *et al*. A novel in-line NIR spectroscopy application for the monitoring of tablet film coating in an industrial scale process [J]. *Talanta*, 2012, 92: 36-37.
- [25] Moes J J, Ruijken M M, Gout E, *et al*. Application of process analytical technology in tablet process development using NIR spectroscopy: Blend uniformity, content uniformity and coating thickness measurements [J]. *Int J Pharm*, 2008, 357(1/2): 108-118.
- [26] Lee M J, Park C R, Kim A Y, *et al*. Dynamic calibration for the in-line NIR monitoring of film thickness of pharmaceutical tablets processed in a fluid-bed coater [J]. *J Pharm Sci*, 2010, 99(1): 325-335.